Review of "A Generalization of the TRIX trophic index to the Adriatic Sea basin" (by Dr. M. Fiori and co-authors (Manuscript under review for journal Nat. Hazards Earth Syst. Sci. doi:10.5194/nhess-2016-69, 2016)

The paper "A Generalization of the TRIX trophic index to the Adriatic Sea basin" compares TRIX index values from in situ data with those from a coupled physics and biogeochemical numerical model implemented in the overall Adriatic Sea. The study aims to demonstrate the compatibility of the model with the in situ TRIX and the necessity to have time series longer than 10 years to evaluate properly the scaling parameters.

The first scope of this paper is very interesting and I think that estimating the trophic index from model outputs could be very useful in supporting policy. However, it should be stated here that the authors are using only one model with a single realization. This is fine with the scope of this study, but it is a limitation for the evaluation of model uncertainty.

Regarding the second object of this study, the demonstration that a sample longer than 10 years is needed for a good estimation of the TRIX scaling parameters is an obvious concept in statistics and in particular in Extreme Values Theory (EVT, see Coles 2001).

By definition the TRIX scaling parameters are yearly (or seasonal) maxima and minima in a period of a fixed length (e.g. 10, 20, 30... years). These parameters follow a Genaralized Extreme Values (GEV) distribution (See Coles et al 2001). A robust estimation of the GEV parameters is based on large-sample theory (asymptotic property), meaning that the times series of the extremes has to be as longer as possible to have stable estimates across periods with different lengths.

As a conclusion, this manuscript can be published after major revision. It is of great interest regarding the comparison between model outputs and observations, however, it has to be revised in all the sections aiming to demonstrate the obvious conclusion that a longer time series provide more robust index estimates. Please, see further suggestions below.

Specific comments

- l.1 I think that the title should be revised accordingly to comments provided above. A possible title could be: "Observed and simulated TRIX trophic index values to the Adriatic Sea basin"

- l.25 The main scope of the MWFD is to achieve Good Environmental Status (GES) of the EU's marine waters by 2020" Please, rephrase this

sentence by reporting the main scope of the Marine strategy and the importance of environmental composite indices in assessing the Good Environmental Status (GES).

- l.29-31 it seems that the authors have modified the previous TRIX definition by defining the TRIX parameters in a different way. As already explained above this is not the case of this study. The authors have only calculated the TRIX parameters with a record length longer than 10 years. Please, remove this sentence or rephrase it accordingly.

- l.36 I would suggest to replace "demonstrates" with "shows"

- l.37 please make this sentence more general. I would suggest something like this: "and as the length of the time series is relevant to get robust index estimates.

- l.54 replace "hence indicators" with "hence composite indicators"

- l.73 this sentence is unclear to me. As the authors reported above the TRIX index has been defined in Volleweider et al. 1998.

- l.74-l77 Extending the TRIX calculation from a period of 10 to one of 20 years does not mean to review the methodology. Please review accordingly.

- l.78-l87 As explained above I do not think this paper present a new TRIX index. Accordingly with Vollenweider et al 1998 the parameter L and M of the TRIX index are defined in log units. The limits for each log transformed variable composing the TRIX index are fixed a priori (see Vollenweider et al Table 3). These limits are fixed a priori from a standard normal distribution of log-transformed variables, being independent from the length of the time series. Of course the correspondence between the fixed limits in log units reported in Vollenweider et al 1998 (see Table 3) and the physical value of the variable will change with the length of the data record. In Vollenweider et al the authors have used a time series of 12 years (1982-1993) because a dataset with longer record length was not available. However, they never wrote in their manuscript that the TRIX limits parameters are fixed with respect a time series of 10 or 12 years. I strongly suggest to review this paper accordingly to this point and to specify that the aims is not to define/modify the TRIX, but to apply the TRIX to observed and model data. This could be very useful for policy makers in order to adopt new strategies.

- l.121-l.125 please review this para accordingly to my comments above (e.g. see l.78-l.87)

- l.180-l.188 Please, specify which correlation coefficient has been used here. I think it is Pearson because log transformed data follow a normal distribution, however it would be better to report this information. Moreover a level of significance (p-value) for correlation should be reported.

- l.235-l.238 Please review this sentence accordingly to comments above (e.g. see l.78-l.87). We need a very large sample size in order to get robust estimates. Recently Sippel et al GRL have shown as μ and σ estimates could be biased by using a time series of 30-year. However this is not the case of the TRIX parameters, as explained above they are fixed a priori on log transformed and standardized variables. To simplify calculation, the ranges have been standardized to 3 log units (see Vollenweider et al 1998).

- Discussion and Conclusions: please review this section by following my comments above. I would focus on the most interesting part of this manuscript which is comparison between simulated and observed TRIX. Moreover, please add somewhere in the text a sentence specifying that in this study only one model with a single realization has been used, limiting the possibility to estimate model uncertainty.

- Suggestions: what about to use a reference period of 30-year to define the TRIX parameters as done for climate extreme indices ($http : //etccdi.pacificclimate.org/list_27_indices.shtml$)?
  A reference period of at least 30-year is suggested by the World Meteorological Organization in order to have robustness of the indices with respect to the length of the time series.
  As an example following Russo et al 2015 using a non-parametric approach to standardize hot days in a heatwave, the authors could define lower and upper TRIX limits on 30-year time series as the 25th and 75th percentile, respectively.

Refrences:

Coles, S. (2001), An Introduction to Statistical Modeling of Extreme Values, 208 pp., *Springer, Berlin.*

Sippel, S., J. Zscheischler, M. Heimann, F. E. L. Otto, J. Peters, and M. D. Mahecha (2015), Quantifying changes in climate variability and extremes: Pitfalls and their overcoming, Geophys. Res. Lett., 42, 99909998.

Russo S, Sillmann J and Fischer E M 2015 Top ten European heat-waves since 1950 and their occurrence in the coming decades Environ. Res. Lett. 10 124003